

Exploring Folding Pathways of Single Proteins Using Mechanical Manipulation

Pétur O. Heidarsson^a and Ciro Cecconi^b

^a*Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen N, Denmark*

^b*CNR-Nano Institute, Department of Physics, University of Modena and Reggio Emilia, Via Guiseppe Camp 213/A, 41125 Modena, Italy*

Abstract. Protein folding is still a major area of active research. Despite significant progress in understanding the underlying principles, we still cannot efficiently predict the folding mechanism for even a moderately sized protein. Proteins are generally thought to fold by diffusion over a three-dimensional energy landscape. Traditional bulk methods have proven to be very powerful in the study of the folding process but they often suffer from inherent ensemble averaging. Single molecule techniques open up new vistas for studying protein folding, allowing direct analysis of the distribution of events that characterize the heterogeneous folding process. Recently it has become possible to directly manipulate individual proteins using optical tweezers. Here we illustrate the experimental strategy and how this approach has provided a fresh perspective on the protein folding problem.

Keywords: Protein folding, single-molecule, optical tweezers, energy landscape

PACS: 82.37.Rs, 87.80.Cc, 87.15.hm

INTRODUCTION

The folding of a protein molecule into its functional three-dimensional structure is a process that has been the focus of active research for decades. It involves a largely unstructured polypeptide of tens to thousands of amino acids spontaneously rearranging into a native structure that is only marginally more stable than the random coil at physiological conditions. Levinthal's paradox states that if the folding process was merely a random search through all of the conformational space available to a protein, then it would take a moderately sized protein (100 amino acids, small in the context of most proteins) longer than the age of the universe to find its native state [1]. Despite the staggering amount of conformations available to a polypeptide of average size, proteins manage to fold in the timescale of milliseconds to seconds. The current view of protein folding is represented in an energy landscape on which the protein diffuses, balancing the favorable reduction in potential energy with the counteracting decrease in conformational entropy as the protein approaches the native state [1]. The energy landscape can have multiple parallel pathways which may or may not include intermediate states (Fig. 1). The energy landscape theory has been successful in explaining biomolecular behavior yet our understanding of the folding process is still incomplete. This is especially true for multi-domain proteins, which represent 75% of the eukaryotic proteome [2], as most folding studies have focused on isolated domains. Traditional bulk studies on protein folding are complicated by the heterogeneity of the process as the protein molecules in a given sample can quickly be de-phased due to the multi-dimensionality of the energy landscape. In addition, rare folding or misfolding events are difficult to examine experimentally. Misfolded proteins are of significant interest as these can accumulate in cells and can lead to diseases such as Parkinson's and Alzheimer's [3].

Single molecule methods have the prospect of looking at the folding process in unprecedented detail and resolving different pathways. Recently it has become possible to mechanically manipulate single protein molecules using optical tweezers [4]. By direct manipulation their behavior under applied force can be described using a convenient reaction coordinate, the molecular end-to-end distance. This method allows a detailed focus on pathways and a reconstruction of the free energy landscape. The single-molecule nature of the experiments allows one to look at distributions rather than averages and, in theory, it should be possible to distinguish rare folding or misfolding events.

In this paper we aim to give a brief overview of optical trapping and its application to the study of protein folding. We describe how molecular constructs are produced for use in these experiments. We briefly illustrate the experimental strategy employed to manipulate individual proteins, and different modes of operation of optical tweezers. Finally, we discuss some case studies.

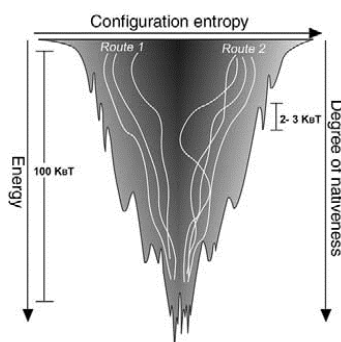


FIGURE 1. Protein folding and energy landscape theory. A protein diffuses through a funnel like energy landscape on its way to the low energy native state. Here are shown two main routes though in principle they can be infinite. Reprinted from [5] with permission from Elsevier.

OPTICAL TRAPPING

Laser tweezers were invented by Ashkin and co-workers in 1986 when they discovered that small particles of high refractive index, such as plastic beads and oil droplets, were attracted to the focal point of a laser beam focused by an objective [6]. This technique, which uses light pressure to trap tiny objects, can be used to manipulate a diverse range of small objects, from living cells down to single atoms, with many applications in biophysics and related fields [4, 7]. Since the invention of the first device, a diversity of laser tweezers set-ups and experimental strategies have been devised and used to study a variety of biological processes. In some cases, the optical trap is created by focusing a single laser beam with a high numerical aperture objective [4]. In other cases, two counter propagating laser beams are used. In these latter devices, the forces applied on the molecule under study can be determined directly by measuring the change in momentum flux of the light beams leaving the trap [8]. Individual molecules cannot be optically trapped and manipulated directly. They need to be attached to movable solid supports. In the following paragraphs, we will describe the experimental strategy used to manipulate individual proteins with optical tweezers.

PROTEIN-DNA CHIMERIC CONSTRUCTS

In order to study protein folding using optical tweezers a chimeric construct is used consisting of a single protein flanked by two ~500 bp DNA molecules that act as molecular handles (Fig. 2). The handles maintain a sufficient distance between the tethering surfaces to avoid unspecific interactions. This method requires proteins that have exposed cysteine residues. Thus, any protein that is amenable to genetic engineering can be studied with this technique. By varying the position of the cysteines, a large set of different pulling axes can be used to study the anisotropy of the folding free energy landscape. The DNA protein coupling reaction is mediated by 2,2'-dithiodipyridine (DTDP). First, an excess of DTDP is allowed to react with the protein to activate the cysteines. This reaction is usually completed within minutes and can be followed spectrophotometrically [9]. The activated proteins are then reacted with a mixture of the two handles carrying a thiol group on one end that can bind to the activated cysteine side chains. This reaction usually takes several hours to produce a good yield of protein-DNA chimeric constructs and can be monitored spectrophotometrically, by atomic force microscopy or by gel electrophoresis. DNA handles have been shown to have negligible effect on the fold and stability of the proteins [10]. Moreover, double stranded DNA (dsDNA) is a convenient molecule to serve as handle in these experiments because its mechanical properties have been extensively studied [11]. At 65 piconewton (pN) dsDNA undergoes a characteristic overstretching transition. This transition can be used to confirm the single molecule nature of the experiment. When two or more DNA molecules are manipulated, they share the applied load and display the overstretching transition at higher force than 65 pN.

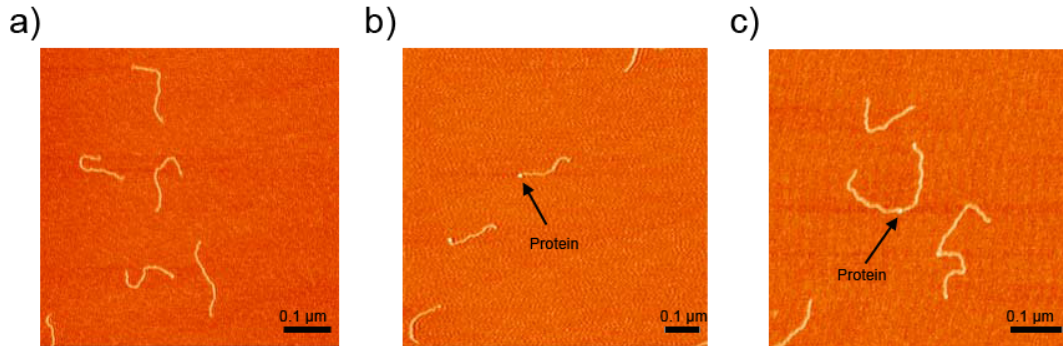


FIGURE 2. Protein-DNA chimeric constructs imaged by atomic force microscopy. a) Free DNA handles, b) ribonuclease H attached to a single handle, and c) ribonuclease H attached to two handles. Reprinted from [9] with kind permission from Springer Science + Business Media: European Journal of Biophysics.

SINGLE-MOLECULE MANIPULATION

The experimental strategy is illustrated in Fig. 3. A protein-DNA molecular construct is manipulated between two polystyrene beads. One bead is held at the end of a pipette by suction while the other is held in the optical trap. The pipette can be moved mechanically with a piezoelectric actuator relative to the trap to stretch and relax the protein. The applied force is determined directly by measuring the change in momentum flux of the light beams leaving the optical trap, while the end-to-end distance of the molecule is estimated using a “light lever” system [8]. These experiments can be performed in a variety of buffer conditions, in the presence of co-solvents or ligands, and at different temperatures [12]. Typically, the molecule under study is manipulated by changing the applied force continuously (force-ramp method). However, the folding mechanisms of a protein can also be studied through experiments in which the applied force is kept constant by a feedback mechanism and fluctuations of the molecule’s extension are monitored over time (force-clamp and force-jump methods).

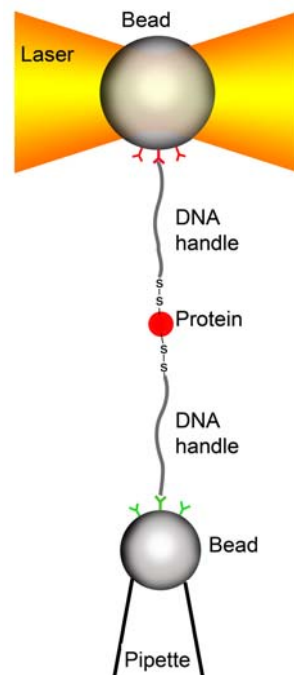


FIGURE 3. Experimental setup (drawing not to scale). The protein is tethered to two polystyrene beads by means of DNA handles. One end of each handle is attached to the protein through a disulfide bond, while the other end is attached to one bead

through either a streptavidin-biotin interaction or an antibody-digoxigenin interaction. The pipette can be mechanically moved relative to the trap while the displacement of the bead in the trap can be used to infer the forces acting on it.

Force-Ramp Experiments

In a typical force-ramp experiment, the molecule is stretched and relaxed multiple times by moving the pipette relative to the optical trap to generate force-extension curves like the one shown in Fig. 4a. When the pipette is moved away from the trap, the force applied on the protein increases until the molecule unfolds. When this happens, the end-to-end distance of the molecule suddenly increases, giving rise to a sharp transition in the stretching trace. Indeed, the extension of an unfolded polypeptide chain is much larger (usually $\sim 20\text{-}30$ nm) than that of the folded structure (usually $\sim 3\text{-}4$ nm). When the pipette is moved toward the optical trap, the molecule is relaxed until it refolds, regaining its original length. When DNA alone is manipulated, force increases and decreases monotonically, and no sharp transitions are observed [10]. Force *vs.* extension curves can be interpolated and analyzed with statistical mechanical models, such as the worm-like chain model (10), to estimate the number of amino acids involved in the observed transitions. This analysis helps identify possible intermediate states along the (un)folding trajectories of a protein [10]. Force-ramp experiments are carried out not only to measure the forces that hold together tertiary and secondary molecular structures, but also to determine kinetics parameters of the unfolding and refolding reactions. Specifically, a molecule can be stretched and relaxed multiple times to generate unfolding/refolding force distributions that can be used to estimate the (un)folding rate coefficients at zero force and the position of the transition state(s) along the pulling axis [10].

Force-Clamp Experiments

The unfolding/refolding processes of a single protein can also be studied through experiments in which the force is kept constant with a feed-back mechanism (force-clamp) and the extension of the molecule is monitored over time. Figure 4b shows extension *vs.* time traces acquired with ribonuclease H (RNase H). In these recordings, RNase H can be observed to hop between a folded and an unfolded conformation. The equilibrium between the two molecular states can be affected by force; higher forces move the equilibrium toward the open state, while low forces favor the closed state. In these experiments, rate coefficients can be obtained directly from the dwell time distributions of the open or closed states, and free energies can be calculated from the ratio of the kinetics coefficients. Force-clamp experiments are ideal to characterize the kinetics and thermodynamics of the interconversion between different molecular conformations at equilibrium. However, these studies are effective only when the rates of the forward and reverse reactions are high enough to allow the observation of a large number of transitions in a relative short amount of time. If the acquisition of a statistically significant number of events require hours of recording, the reliability of the measurements can be compromised by the mechanical and optical instability of the instrument. In these cases, the folding of a protein can be more effectively studied through force-jump experiments.

Force-Jump Experiments

In a force-jump experiment the force is increased (jumped) or decreased (dropped) quickly to a preset force value and kept constant with a feedback mechanism until an unfolding or refolding event is observed (Fig. 4c). These experiments allow the direct measurement of rate constants in force ranges where the probability of observing either an unfolding or refolding event is high. The force-jump/force-drop cycle can be repeated multiple times until a large number of events have been observed. At each cycle, the time that the molecule takes to unfold or refold can be directly measured and plotted in dwell time histograms to estimate the rate coefficients. From the dependence of the rate coefficients on the applied force, kinetics parameters, such as position of the transition state along the reaction coordinate and rates at zero force, can be estimated. This method was introduced for the first time to study the kinetics and thermodynamics of the (un)folding mechanism of an RNA molecule [13].

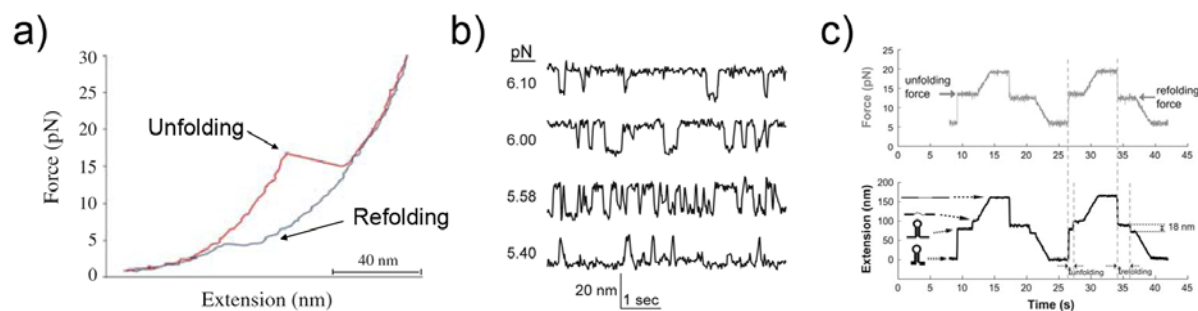


FIGURE 4. Data acquired through different modes of operation of optical tweezers. a) Force-extension cycle of RNase H from a force-ramp experiment. The stretching trace is characterized by a sharp discontinuity at ~ 16 pN that marks the mechanical unfolding of the molecule. The relaxation trace is characterized by a refolding transition at ~ 5 pN. b) Extension vs. time traces of RNase H at various forces in a force-clamp experiment. The probability of populating either the unfolded conformation (upper level) or the folded conformation (lower level) is affected by the applied force. c) Two successive force-jump/force-drop cycles of TAR RNA. The top and bottom panels show the force and extension, respectively, monitored over time. Figure 4a is reprinted from [9] with kind permission from Springer Science + Business Media: European Journal of Biophysics. Figure 4b is reprinted from [10] with permission from AAAS. Figure 4c is reprinted from [13] with permission from Elsevier.

CASE STUDIES

A number of systems have been studied with the methods described above. Impressive results have emerged from these studies that have furthered our understanding of protein folding. A subset of these results is described below.

Ribonuclease H

Optical tweezers were employed to study the folding of a single globular protein for the first time by Cecconi and co-authors [10]. By manipulating one molecule at a time using DNA molecular handles, these authors characterized the (un)folding processes of individual RNase H molecules with unprecedented details, uncovering information inaccessible to more traditional bulk techniques. Bulk studies had previously revealed the existence of a partially structured intermediate (*I*) in the folding trajectories of RNase H [14]. The intermediate, however, formed within the dead time of the measuring instrument (12 ms) and thus could not be observed and characterized directly. By manipulating one molecule at a time and slowing down the refolding process with tension, Cecconi and co-authors were instead able to watch directly single molecules fold part way to the intermediate structure, and then go from there to the folded protein. Using this experimental approach, they showed that RNase H folds through a compact, thermodynamically distinct intermediate state, which is both obligatory and on-pathway to the native structure. In force-clamp experiments, they observed the protein fluctuate between the intermediate and the unfolded state (*U*), and characterized the *U* to *I* transitions in terms of kinetic and thermodynamic parameters. They also characterized the mechanical properties of the intermediate and showed that it is highly pliant and stretchy under force. This is a property characteristic of structures, such as those of molten globule-like states, mainly stabilized by secondary interactions and weak, non-stereospecific tertiary contacts.

T4 Lysozyme

In optical tweezers experiments the denaturant (force) can be selectively applied to specific portions of a protein. This is a unique capability of optical tweezers that makes this technique particularly well suited to study the cooperativity of the folding processes of different protein domains. An example of the effectiveness of these single molecule studies to investigate the cooperative communication between different structural regions of a protein is represented by the work published in *Nature* by Shank *et al.* [15]. In this work, the authors studied the folding cooperativity of T4 lysozyme. They applied force selectively to either the N-terminal or C-terminal domain of this protein and used the Crooks fluctuation theorem to probe the status of the protein portion not directly subject to force. The results of these studies showed that the folding processes of the two structural domains of T4 lysozyme are strictly coupled. Additionally, by creating a topological variant with a circular permutation, Shank *et al.* showed that helix A is a key determinant of the cooperative unfolding of T4-lysozyme.

Calmodulin

Stigler *et al.* studied the calcium sensor calmodulin, a two-domain EF-hand protein that is involved in a number of signaling cascades [16]. Using ultra-stable high-resolution optical tweezers and performing force-clamp experiments, they observed a total of six intermediate states in the folding mechanism of calmodulin. Four of these intermediate states were on-pathway in a relatively fast folding process. The presence of two off-pathway intermediate states, which had non-native inter-domain interactions, slowed the overall folding process. From their results, Stigler *et al.* were able to show that calmodulin folds through three competing pathways where different intermediates were populated giving rise to different folding rates. Their results lend support to the theory of kinetic partitioning in which a fraction of a given set of molecules folds fast towards the native state while others are kinetically trapped and fold slower. The overall result demonstrates the high degree of complexity of the dynamical folding network of these signaling molecules.

CONCLUSIONS

Here we have described how optical tweezers can be used to study chimeric protein-DNA constructs on the single-molecule level. This method is robust and generally applicable to any protein that can be modified with genetic engineering. A wide range of experimental conditions can be applied and experiments can be chosen that fit the specific system under study.

The advent of single molecule methods has opened up new vistas of possibilities when it comes to exploring the phenomenon of protein folding. Optical tweezers permit a direct and gentle manipulation of biomolecules within both the thermodynamic and kinetic regimes of the energy landscape of a protein. Here, we have outlined a subset of the vastly growing number of studies in which optical tweezers have been used to study the folding behavior of proteins. As the technology of optical trapping advances further we can expect exciting results on even larger and more complex systems in the future.

REFERENCES

1. K. A. Dill and H. S. Chan, *Nat Struct Biol* **4** (1), 10-19 (1997).
2. M. B. Borgia, A. Borgia, R. B. Best, A. Steward, D. Nettels, B. Wunderlich, B. Schuler and J. Clarke, *Nature* **474** (7353), 662-665 (2011).
3. J. H. Han, S. Batey, A. A. Nickson, S. A. Teichmann and J. Clarke, *Nat Rev Mol Cell Biol* **8** (4), 319-330 (2007).
4. J. R. Moffitt, Y. R. Chemla, S. B. Smith and C. Bustamante, *Annu Rev Biochem* **77**, 205-228 (2008).
5. J. N. Onuchic and P. G. Wolynes, *Curr Opin Struct Biol* **14** (1), 70-75 (2004).
6. A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm and S. Chu, *Opt Lett* **11** (5), 288 (1986).
7. D. G. Grier, *Nature* **424** (6950), 810-816 (2003).
8. S. B. Smith, Y. Cui and C. Bustamante, *Methods Enzymol* **361**, 134-162 (2003).
9. C. Cecconi, E. A. Shank, F. W. Dahlquist, S. Marqusee and C. Bustamante, *Eur Biophys J* **37** (6), 729-738 (2008).
10. C. Cecconi, E. A. Shank, C. Bustamante and S. Marqusee, *Science* **309** (5743), 2057-2060 (2005).
11. S. B. Smith, Y. Cui and C. Bustamante, *Science* **271** (5250), 795-799 (1996).
12. H. Mao, J. R. Arias-Gonzalez, S. B. Smith, I. Tinoco, Jr. and C. Bustamante, *Biophys J* **89** (2), 1308-1316 (2005).
13. P. T. Li, D. Collin, S. B. Smith, C. Bustamante and I. Tinoco, Jr., *Biophys J* **90** (1), 250-260 (2006).
14. T. M. Raschke and S. Marqusee, *Nat Struct Biol* **4** (4), 298-304 (1997).
15. E. A. Shank, C. Cecconi, J. W. Dill, S. Marqusee and C. Bustamante, *Nature* **465** (7298), 637-640 (2010).
16. J. Stigler, F. Ziegler, A. Gieseke, J. C. Gebhardt and M. Rief, *Science* **334** (6055), 512-516 (2011).